REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1995	Final Report (Mar 93 - Apr 95)

4. TITLE AND SUBTITLE

Efficacy of repeated psychophysiological detection of deception testing

5. FUNDING NUMBERS

DoDPI93-P-0025

6. AUTHOR(S)

Andrew B. Dollins and Victor L. Cestaro

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Defense
Polygraph Institute
Building 3165
Fort McClellan, AL 36205-5114

8. PERFORMING ORGANIZATION REPORT NUMBER

DoDPI94-R-0013

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Department of Defense
Polygraph Institute
Building 3165
Fort McClellan, AL 36205-5114

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DoDPI94-R-0013

DoDPI93-P-0025

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Public release, distribution unlimited.

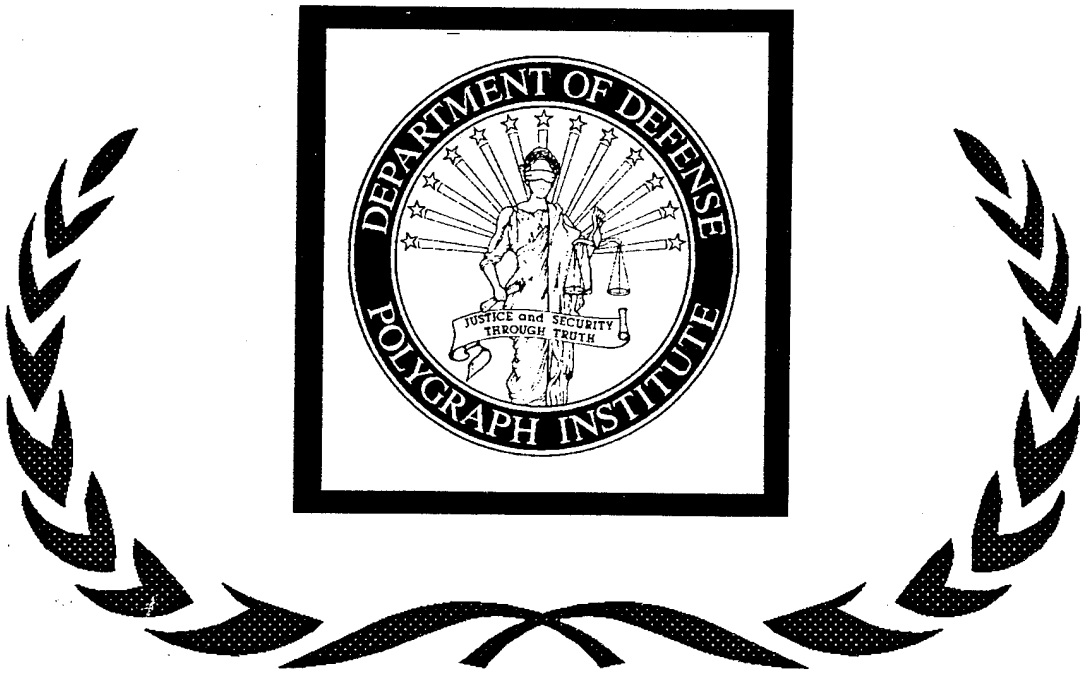
12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Physiological measures were recorded during repeated psychophysiological detection of deception (PDD) tests to determine if physiologic response levels change with test repetition. Two groups of 22 healthy male subjects completed six Peak of Tension PDD tests on each of two test days. A minimum between test day interval of six days was maintained. The treatment group was programmed to respond deceptively to one of seven test questions while the control group was programmed to respond truthfully to all questions. The respiration and Galvanic Skin Response (GSR) line lengths, GSR peak response amplitude and latency, and cardiovascular inter-beat-interval (IBI) were calculated for each response. Analyses indicated that: except for GSR peak response latency, differential physiological reactivity during a PDD test did not change significantly during repeated tests or days; there was a decrease in average respiration line lengths during the beginning test(s) of each day; and, differential changes in average respiration line length, GSR peak latency, and cardiovascular IBI responses corresponded to deception. Power analyses are presented to assist in result interpretation. It is suggested that PDD decision accuracy, concerning subject veracity, should not decrease during repeated testing.

14. SUBJECT TERMS	15. NUMBER OF PAGES
psychophysiological detection of deception, peak of tension, repeated measures, respiration, galvanic skin response, heart rate, power analysis	49
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	



Efficacy of Repeated Psychophysiological Detection of Deception Testing

Andrew B. Dollins, Ph.D., Victor L. Cestaro, Ph.D.,
and Donald J. Pettit, Ed.D.

April 1995

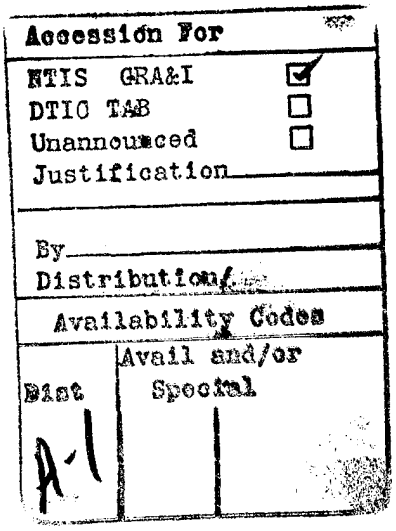
Department of Defense Polygraph Institute
Fort McClellan, Alabama 36205-5114
Telephone: 205-848-3803
FAX: 205-848-5332

Report No. DoDPI94-R-0013

Efficacy of Repeated Psychophysiological Detection of Deception Testing

Andrew B. Dollins, Ph.D., Victor L. Cestaro, Ph.D.,
and Donald J. Pettit, Ed.D.

April 1995



Department of Defense Polygraph Institute
Ft. McClellan, AL 36205

Director's Foreword

This study addresses the important issue of reliability as it relates to psychophysiological detection of deception (PDD) in a manner never before undertaken. Most published studies regarding PDD reliability report inter-rater and intra-rater scoring comparisons and thus are not really reliability findings. Reliability results, as reported in this study, are based on physiologic data collected from the same individual, while responding to the same test questions, on two separate occasions.

What makes this study unique is that the response measures were quantified and then compared, whereas other studies, involving test retest on the same subject, merely compared diagnostic accuracies of the results of each test. The finding in this study, that the response patterns did not change significantly between test and retest, provides additional, but not definitive, information to support the practice of repeated testing on individuals under criminal and screening circumstances.

The fact that any type of test can be reliable without being valid, but can't be valid without being reliable, suggests that every PDD test that is taught at the Department of Defense Polygraph Institute (DODPI) should have validity and test retest reliability data included in the administration manuals. Only one of the PDD tests taught at DODPI, the Zone Comparison Test (ZCT) has been studied under test retest reliability conditions. Test retest reliability studies will be pursued and will include all testing formats currently taught at the Institute.

William J. Yankee, Ph.D.
Director

Acknowledgments

The authors wish to express special thanks to: the volunteers who participated in the study; Charlene L. Stephens, Master Sergeant Randall S. Reynolds, United States Air Force, and the support staff at the Department of Defense Polygraph Institute (DoDPI), without whom this study could not have been completed. The authors would also like to thank Chief Warrant Officer Four Madison L. Mumbauer, United States Army, for his editorial assistance, and the Director of the DoDPI, Dr. William J. Yankee, for his continued support throughout the project. This study was supported by funds from the DoDPI as project DoDPI93-P-0025. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

Abstract

DOLLINS, A. B., CESTARO, V. L., and PETTIT, D. J. Efficacy of repeated psychophysiological detection of deception testing, April 1995, Report No. DoDPI94-R-0013. Department of Defense Polygraph Institute, Ft. McClellan, AL 36205. -- Physiological measures were recorded during repeated psychophysiological detection of deception (PDD) tests to determine if physiologic response levels change with test repetition. Two groups of 22 healthy male subjects completed six Peak of Tension PDD tests on each of two test days. A minimum between test day interval of six days was maintained. The treatment group was programmed to respond deceptively to one of seven test questions while the control group was programmed to respond truthfully to all questions. The respiration and Galvanic Skin Response (GSR) line lengths, GSR peak response amplitude and latency, and cardiovascular inter-beat-interval (IBI) were calculated for each response. Analyses indicated that: except for GSR peak response latency, differential physiological reactivity during a PDD test did not change significantly during repeated tests or days; there was a decrease in average respiration line lengths during the beginning test(s) of each day; and, differential changes in average respiration line length, GSR peak latency, and cardiovascular IBI responses corresponded to deception. Power analyses are presented to assist in result interpretation. It is suggested that PDD decision accuracy, concerning subject veracity, should not decrease during repeated testing. It is further suggested that pneumograph line length and cardiovascular IBI are reliable response measures which may be sensitive to physiological changes associated with deception.

Key Words: psychophysiological detection of deception (PDD), peak of tension (POT), repeated measures, respiration, galvanic skin response (GSR), heart rate, power analysis

Executive Summary

DOLLINS, A. B., CESTARO, V. L., and PETTIT, D. J. Efficacy of repeated psychophysiological detection of deception testing, April 1995, Report No. DoDPI94-R-0013. Department of Defense Polygraph Institute, Ft. McClellan, AL 36205.

Physiological measures were recorded during repeated psychophysiological detection of deception (PDD) tests to determine if physiologic response levels change with test repetition. Two groups of 22 healthy male subjects completed six Peak of Tension PDD tests on each of two test days. A minimum between test day interval of six days was maintained. Question presentation parameters were held constant through the use of digital recordings. All subjects were asked the same series of questions in the same order throughout each test. The treatment group was programmed to respond deceptively to one of seven test questions while the control group was programmed to respond truthfully to all questions. The Lafayette Factfinder polygraph, which was used throughout testing, was modified to permit digital recording of response activity which was independent of operator sensitivity and offset adjustments. The respiration and Galvanic Skin Response (GSR) line lengths, GSR peak response amplitude and latency, and cardiovascular inter-beat-interval (IBI) were calculated for each response.

Analyses indicated that: except for GSR peak response latency, differential physiological reactivity within a PDD test does not change significantly during repeated tests or days; there was a decrease in average respiration line lengths during the beginning test(s) of each day; and, changes in average respiration line length, GSR peak latency, and cardiovascular IBI responses corresponded to deception. The results of statistical power analyses, which determines the probability that the null hypothesis (i.e., no significant differences between measures) is correctly evaluated, are presented to assist in result interpretation. It is suggested that PDD decision accuracy, concerning subject veracity, should not decrease during repeated testing. It is further suggested that pneumograph line length and cardiovascular IBI are reliable response measures which may be sensitive to physiological changes associated with deception.

Table of Contents

List of Figures

The United States Department of Defense, various law enforcement agencies, and officers of the court routinely use a psychophysiological detection of deception (PDD) examination to determine an individual's truthfulness concerning topics of interest (Office of Technology Assessment, 1983, pp. 1-8; Lykken, 1981, pp. 1-4). The theory underlying PDD is that physiologic reactivity, in response to the presentation of a stimulus, varies with the personal relevance of the stimulus and, more so, with attempts to conceal that relevance. The typical PDD examination is designed to elicit physiologic reactions from the examinee in response to questions concerning topic(s) of interest. Variability in Galvanic Skin Resistance (GSR), respiratory rate and/or volume, and heart rate/blood pressure are typically assessed (visually) by PDD examiners in the field. An increase in reactivity, defined as a change in response rate and/or amplitude, is interpreted as indicative of the examinee's truthfulness regarding the questions of interest.

Numerous valid criticisms have been expressed regarding the PDD process and associated assumptions (Furedy, 1986; Lykken, 1981; Office of Technology Assessment, 1983, pp. 29-43). Among those are the criticisms of validity and reliability of results. Validity is defined (Campbell, 1989, p. 749) as the degree to which a test measures what it is supposed to measure. Validity of a PDD examination would be measured as the degree of agreement between examiner decisions and ground truth (facts). Virtually all PDD studies attempt to assess the validity of PDD by comparing examiner decisions to ground truth. Definitions of ground truth range from experimental programming (i.e., asking subjects to participate in mock crimes so guilt and innocence are known quantities; Barland & Raskin, 1975) to decisions made by panels of experts who have reviewed case reports (Bersh, 1969). While questions of validity are very important, they are moot if reliable examination results are not obtained. Reliability is defined (Campbell, 1989, p. 629) as the degree to which a test measures the same thing consistently. A test of PDD examination reliability would require testing the same individual twice, using the same procedures. If PDD examinee responses are not consistent (among and/or between different measures), it is unlikely that questions of validity can ever be properly addressed. There have been numerous studies of interexaminer reliability in evaluating physiological data collected during a PDD examination (e.g., Horvath, 1977; Horvath & Reid, 1971; Hunter & Ash, 1973; Slowick & Buckley, 1975). Such studies are important in that they examine the consistency of data interpretation among examiners. These studies have not, however, investigated the reliability of physiologic responses.

Few studies report results concerning the consistency of examinee responses. An exploratory study completed by Ellson, Davis, Saltzman, and Burke (1952) was designed to examine the

GSR responses of 10 male subjects using a variation of what is now labeled a stimulation card test (Abrams, 1989, pp. 120-122). They conclude that "one repetition of the detection procedure does not noticeably affect the success of the GSR as an indicator" of deception (Ellson et al., 1952, p. 7), but refer to no inferential statistics. Results of a second, similar, study confirm this hypothesis "unless the subject is told that the first attempt was successful" (Ellson et al., 1952, p. 11). Lieblich, Naftall, Shmueli, and Kugelmass (1974) employed a similar stimulation paradigm and GSR measure. They report that identification of deception was improved by repeating the same question sequence 10 times. Balloun and Holmes (1979) recorded the responses of 16 male subjects during two 5-question PDD examinations, separated by 30 seconds, administered using the Guilty Knowledge Questioning Technique (Lykken, 1960). They found that responses were attenuated during the second administration of the test and suggest that repeated examinations may be invalid. Grimsley and Yankee (1986) employed the Relevant/Irrelevant Question Technique to examine 80 male and female subjects on three occasions (separated by 24 hours). They found a non-significant decrease in accuracy between examinations 1 and 2, but no difference in accuracy between examinations 1 and 3. They conclude that overall accuracy rates are increased by evaluating multiple examinations. Yankee (1993) used the Control Question Technique (Reid, 1947) and a somewhat more realistic paradigm to investigate the accuracy of repeated examinations. Subjects (N = 72) were examined on two occasions, separated by 24 hours. Half of the subjects were programmed guilty via participation in a mock crime. Yankee also reported a decline in accuracy, though smaller in magnitude than that reported by Balloun and Holmes (1979), between the two examinations.

None of the investigations of repeated PDD examinations report data quantification beyond visual examination of physiological data. Decisions were usually based on visual inspection alone. Accurate absolute response levels are not mentioned. The very fundamental question of whether absolute response level differences occur during repeated examinations has not been investigated. The effect of a moderate delay between repeated examinations has also not been examined. The effect of such delays is simply not known and field examiners must rely on anecdotal knowledge for guidance. The current study is designed to examine response levels throughout repeated PDD examinations. A relatively simple variation of the Peak of Tension paradigm was chosen under the assumption that the results would generalize to more complex paradigms which use questions of greater personal relevance.

Method

Subjects

Forty-four, native English speaking, healthy males [mean age (SD) = 29.2 (7.8) years; range = 19 to 47] participated in this study. They were military personnel or Department of the Army civilian employees and were not paid for their participation. Thirty-nine of the subjects had never participated in a PDD examination before. The remaining five had not participated in a PDD examination within the last two years. Thirty-five of the subjects reported themselves to be medication free. The remainder had ingested pain/relaxant (3), anti-inflammatory (1), antibiotic (2), and antihistamine (3) medication within the 12 hour period prior to the examination. Females were not included because of possible variations in GSR (over-time) caused by hormonal secretions associated with the menstrual cycle.

Examiner

All PDD examinations were conducted by the same examiner, who had been trained at the United States Army Polygraph School and was certified by the United States Army as competent to administer PDD examinations. The examiner had administered approximately 500 field examinations during the 5 years prior to the study and was an instructor at the Department of Defense (DoD) Polygraph Institute at the time of the study. The examiner was not aware of whether subjects belonged to the control or treatment groups.

Apparatus

Data were collected using a Lafayette (Lafayette, IN) Factfinder (Model 76740/76741) polygraph equipped with three Cardio | Aux | Pneumo | GSR modules (Model 76477-G), one GSR module (Model 76480-G), and one electronic stimulus marker module (Model 76351-GET). Two of the multifunction modules (Model 76477-G) were used to record respiratory activity by setting the function selector to Pneumo and the third was used to record cardiovascular activity by setting the selector to Cardio-1. A circuit was added to the electronic stimulus marker module to allow control of the marker via signals from a computer RS-232 serial port. Lafayette sensors were used to measure GSR (Model 7664), respiration (Model 76513-1G & 76513-2B), and cardiovascular activity (Model 76530).

An electronic circuit (schematically depicted in Appendix A) was designed and built in-house to amplify voltages from the Lafayette modules used to measure GSR, respiration, and cardiovascular activity. The amplified voltages were not affected by sensitivity or centering adjustments made to the instrument. Connection points for signal acquisition were: 1) GSR module - Pin 1, integrated circuit (I.C.) U1; 2) Cardio | Aux | Pneumo | GSR module - for respiration - Pin 3 of I.C. U1;

and 3) Cardio | Aux | Pneumo | GSR module - for cardiovascular activity - Pin 7 of I.C. U2. The amplification circuit contained potentiometers which could be used to adjust the pre-amplifier voltage offset. A DC offset was indicated to be positive or negative by red/green LEDs mounted near the potentiometers. Amplification gains during testing were set at: 47x for the Pneumo/respiration channels; 10x for the Cardio channel; and 5x for the GSR channel. Post-amplification signals were connected to a female 9-pin D connector. The amplification circuit module was inserted in an empty slot on the Lafayette polygraph and powered by the polygraph's internal power supply. The potentiometer controls, LED voltage indicators, and the 9-pin D connector were user accessible on the surface of the polygraph. Post-amplification physiologic signals were digitized using a Keithley Metrabyte (Taunton, MA) DAS-16F analog-to-digital converter mounted in an IBM PS/Value Point (Armonk, NY) Model 433DX microcomputer. Software was written in-house to digitize the physiologic signals at a rate of 256 samples/second.

A second micro-computer (Model 248, Zenith Data Systems, Chicago, IL), was used for question presentation. The questions used throughout testing were digitized and recorded to computer hard disk using a Sound Blaster board (Model 16ASP, Creative Labs Inc, Milpitas, CA). A parallel port interface (Speech Thing, Covox Inc., Eugene, OR), connected to a Radio Shack (Fort Worth, TX) integrated stereo amplifier (Model SA-155) and two speakers (Model Minimus-77), was used to present the questions. This system ensured that each question was presented with the same inflection, and at the same volume, each time it was repeated. Computer software was written in-house to allow the examiner to present questions and digitize data by moving a cursor on the computer screen. Activation of the question presentation and data acquisition software was achieved via a serial port request-acknowledge algorithm.

Subjects verbal responses were recorded on cassette tape (Tascam Model 134 4-channel recorder, TEAC, Montebello, CA) using a lavaliere microphone (Model 570S, Shure, Evanston, IL) held in place by a cord placed over the examinee's shoulders. The recorder was located in an adjacent room. Excerpt recording was controlled via the software running on the question presentation computer. The question presentation computer serial port and an in-house built interface for the cassette recorder were used for this purpose. Sound features of the audio recordings were extracted and examined as possible indexes of deception, as reported elsewhere (Cestaro & Dollins, 1994).

PDD testing was conducted in a carpeted, 3.5 x 3.66 m partially sound-attenuated room. Each examination was recorded on video tape using two ceiling and one wall-mounted video

cameras. The examination was also monitored through a two-way mirror by a collaborator located in an adjacent room.

Subjects were seated in a Lafayette adjustable-arm subject chair (Model 76871, Lafayette, IN) during PDD testing. The chair was positioned beside and slightly in front of the examiner's desk. This position allowed the examiner to monitor the examinee's movements but not vice versa. The polygraph was mounted in a double pedestal examiner's desk (Lafayette Model # 76183). The question presentation and data acquisition computers and monitors were positioned on a table next to the examiner's desk and out of the examinee's sight during testing. The speakers, through which the questions were played, were located six feet behind, and one foot above, the back of the examinee's chair. The examinee's field of view, throughout testing, contained a wall of uniform color, a stationary video camera, and, above the video camera, a piece of paper with numbers and words written on it. Video cameras were also placed in the ceiling above the examiner's desk and behind the subject.

Procedure

Subjects were randomly assigned to the treatment or control groups, with the constraint that no more than three control or treatment group participants were tested consecutively. Twenty-two subjects were assigned to each group. Each subject participated in two examination sessions. The two examinations were separated by at least six working days. Subjects completed six Peak of Tension PDD tests during each examination session.

Upon arrival at the DoD Polygraph Institute (Fort McClellan, AL), each subject was escorted to a secluded briefing room and asked to read a brief description of the research project (see Appendix B). Subjects who indicated that they would participate were asked to read and sign a volunteer agreement affidavit (see Appendix C). Their questions were then answered. A brief biographical/medical questionnaire was then completed, to ensure that the subject was in good health and not currently taking medication which could interfere with the PDD examination results (see Appendix D). The subject was then asked to complete a number search task, which was referred to as an anagram task. During this task, the participant circled six sequences of a two-digit number which was repeated five consecutive times (in any direction) in a 20 x 30 matrix of two digit numbers. The matrix consisted of numbers between 60 and 69 for the programmed deceptive subjects - who circled the number 64 (see Appendix E), and 80 to 89 for the programmed non-deceptive subjects - who circled the number 84 (see Appendix F). When the anagram task was completed, the subject was asked to write his name and the number he circled on two 7.62 x 12.7 cm cards. One card was retained by an investigator and the second concealed in the subject's pocket. The PDD examination procedure was briefly explained to the subject. It was

emphasized that during the PDD examination the subject should not reveal which number he had circled when completing the anagram task. It was further emphasized that he should make every attempt to remain relaxed, even if he felt himself begin to react (increased heart rate, perspiration on hands, tightening of occlusive cuff) during the examination. The subject was then escorted to the examination room and introduced to the examiner.

The examiner greeted each subject, then reviewed the biographical/medical questionnaire with him to ensure it's accuracy. No other pre-test questions were asked by the examiner. The examiner then briefly explained the sensors, procedures, and theory of PDD (see Appendix G). The examiner explained that the polygraph measured physiological reactions - and not deception per se. It was further explained that the subject's physiological responses were likely to change during deception. It was suggested that fear of detection during deception altered the normal physiological response pattern and that these changes may be evident in the signals recorded during the PDD examination. The examiner described this response as being similar to the fight-or-flight reaction used to describe a fear response during military training. The examiner then reviewed the questions to be asked during data collection, with the subject, by playing the recorded questions.

All questions asked by the subject were then answered. He was then seated in the examination chair and the sensors were attached. Respiration was monitored using convoluted (pneumo) tubes placed around the thoracic area and abdomen. GSR was measured using stainless steel field electrodes placed, without paste, on the volar surface of the distal phalanges of the examinee's right hand index and ring fingers. Cardiovascular activity was monitored using an occlusive cuff placed over the brachial artery of the left arm. The pneumo tube vents were closed and the DC offsets for the pneumo and GSR were adjusted to zero. The sensitivity of these recording channels was then adjusted on the polygraph. Next, the occlusive cuff was inflated to 90 mmHg, massaged to remove wrinkles, then deflated to 48 mmHg. The pressure was then adjusted, as necessary, to achieve a 2 mmHg dial deflection between diastole and systole on the sphygmomanometer. The amplifier DC offset was then adjusted to zero, and polygraph sensitivity adjustments were made.

Each PDD test was composed of the following series of statements and questions, which were presented via computer recorded voice.

X	The test is about to begin.
01	Did you complete an anagram for the number 60?
02	Did you complete an anagram for the number 61?
03	Did you complete an anagram for the number 62?

04 Did you complete an anagram for the number 63?
05 Did you complete an anagram for the number 64?
06 Did you complete an anagram for the number 65?
07 Did you complete an anagram for the number 66?
XX The test is now complete, please continue to sit still while I turn the instrument off.

If the examiner judged the physiological signals recorded on the polygraph chart to contain artifacts, the previous question was repeated. The examiner played the pre-recorded message "please remain still" if he judged that the examinee was producing unnecessary and/or excessive movements. When data collection for each test was completed, the pressure in the occlusive cuff was vented and the subject was instructed to "please relax while I prepare for the next test." If subjects appeared to be sleepy, they were also reminded of the importance of the study and encouraged to remain alert. The next PDD test was begun approximately three minutes later. The occlusive cuff was inflated, as described above, and DC offsets for the GSR and cardiovascular activity amplifiers were adjusted prior to beginning the next test. This process was repeated until six tests were completed, after which the sensors were removed. The subjects were then asked to read and sign a debriefing form (see Appendix H), reminded to return the following week, and escorted out of the building.

Subjects returning for a second test session were escorted to a briefing room. They were reminded of the number circled during the previous session and asked to conceal the second card, indicating the number circled, in a pocket. They were reminded not to reveal the number they had circled to the examiner, then escorted to the examination room. The examiner again reviewed the biographical/medical questionnaire from their previous session to ensure that no significant changes had occurred. Six additional PDD tests were completed, as described above. When the examination was completed, participants were thanked for their cooperation, asked to read and sign a second debriefing form (see Appendix I), and escorted out of the building.

Data Reduction

The upper and lower pneumograph, GSR, and cardiovascular responses to each question were sampled at a rate of 256 samples per second for 14 seconds. Data sampling was initiated by the stimulus marker indicating that playback of the recorded question had ended. The data for each channel were smoothed to remove noise inherent in the instrument and/or amplifier used. Smoothing was implemented by substituting the average of the 50 points pre- and pro-ceeding a data sample (i.e., a running average of 101 data points) for that sample. The first and last 50 data points of each epoch were then omitted from the epoch.

This smoothing procedure was empirically determined to be the optimal solution to reducing noise in the recorded signal.

The data collected during day 1, test 3, questions 61 through 64 were lost, due to experimenter error, for 5 subjects (3 deceptive and 2 non-deceptive). Each response was reviewed for movement artifact contamination by three psychophysiologists who were blind to the treatment condition in which the sample was collected. Responses identified as containing movement artifacts by two or more reviewers were marked as missing data and omitted from further processing. All responses with amplitudes which exceeded the limits of the analog-to-digital converter were marked as missing data.

The following statistics were calculated for the remaining 13.6 second epochs. Line length of the upper and lower pneumograph (Pn1-LnL and Pn2-LnL, respectively), a technique introduced by Timm (1979, 1982a, 1982b), and GSR (GSR-LnL) data were calculated using a between point interval of 0.00390625 (i.e., 1/256). GSR peak amplitude (GSR-Amp) was calculated as the Peak Amplitude minus (0.5 * (Trough 1 + Trough 2) amplitudes). Troughs and peaks were identified as the first point where the subsequent 200 samples were greater (trough) or less (peak) than that point. If a peak was not identified within the first 7 seconds of data sampling, the peak amplitude values for the epoch were set to 0.000. Trough 1 was the first trough occurring prior to the peak or the first data sample if a peak but no trough was located. Trough 2 was the first trough identified after the peak. GSR peak latency (GSR-Ltc) was calculated, in seconds, relative to the first data point collected, for analysis where peaks were found. If a peak was not identified then the peak latency was considered missing data. The average heart rate inter-beat interval (CRD-IBI) epoch was calculated by determining the latency between the first and last R-wave peak found during the 13.6 second epoch and dividing by the total number of peaks found during the epoch - minus one.

The mean and standard deviation of responses recorded under each condition of the independent variables (group, day, test, and question) were calculated and only values within two standard deviations of the mean were retained for further analysis. (Note that data previously described as missing were omitted from this calculation.) All missing data were replaced by means from the appropriate condition combination. The proportion of missing data for each measure - by deceptive/non-deceptive group, respectively, was: Pn1-LnL - .07 / .07; Pn2-LnL - .07 / .09; GSR-LnL - .14 / .10; GSR-Amp - .15 / .12; GSR-Ltc - .25 / .20; and, CRD-IBI - .05 / .07.

It was observed that more than 50% of the GSR line length and amplitude data were missing for 2 subjects in each group

and that more than 50% of the GSR peak latency data were missing for 6 subjects in each group. The data for these subjects were not analyzed for these measures.

Data Analysis

Statistical analyses were calculated using SYSTAT for DOS (Version 5.0) and Windows (Version 5.04 - SYSTAT, Inc., Evanston, IL). The Pn1-LnL, Pn2-LnL, GSR-LnL, GSR-Amp, GSR-Ltc, and CRD-IBI response measures were initially analyzed using a between groups, within subjects 2(between-group) x 2(within-day) x 6(within-test) x 6(within-question) repeated measure analysis of variance (ANOVA). As mentioned above: 22 subjects per group were included in the Pn1-LnL, Pn2-LnL, and CRD-IBI analyses; 20 subjects per group were included in the GSR-LnL and GSR-Amp analyses; and 16 subjects per group were included in the GSR-Ltc analysis. A completely within subjects 2(day) x 6(test) x 6(question) repeated measure ANOVA was subsequently calculated, where appropriate, to resolve group main and interaction effects. The degrees of freedom used in calculating each mean square error term and F statistic were reduced by the proportion of missing data for that measure. F statistic probabilities of repeated measure effects with more than two levels were corrected for violations of sphericity assumptions using the Greenhouse - Geisser (1959) epsilon (ϵ). Orthogonal planned comparisons (Winer, 1971, pp. 172-215) were used to evaluate significant ($p < 0.05$) test and question main effects. The comparisons chosen to evaluate test effects were: (a) test 1 versus tests 2, 3, 4, 5, and 6; (b) test 2 versus tests 3, 4, 5, and 6; (c) test 3 versus 4, 5, and 6; (d) test 4 versus tests 5 and 6; and (e) test 5 versus test 6. Significant question effects were evaluated by comparing the measures recorded in response to questions concerning the numbers 62, 63, 64, 65, and 66 to those recorded in response to the remaining questions. For example, the responses following the question concerning the number 62 were compared to those concerning the numbers 61, 63, 64, 65, and 66.

The statistical power of each ANOVA F-test was calculated to assess the probability that the null hypothesis of no difference between the treatment means would be correctly rejected when the hypothesis was false (Williams & Zimmerman, 1989). Effect sizes were calculated as described by Cohen (1988, pp. 531-535), then converted to the noncentrality parameter, lambda, by multiplying the squared effect size by the number of observations in each analysis (Cohen, 1988, p. 550). It was necessary to convert effect sizes to a noncentrality parameter and calculate power directly rather than use Cohen's (1988) effect size directly because Cohen's (1988) tables underestimate the power of factorial designs (Koele, 1982). The denominator degrees of freedom used in the power calculations were reduced by the percent of missing data, as described above. Because the number of subjects in this design was relatively

large, the power of each main effect and interaction was calculated using Laubscher's (1960, Formula 6) square root approximation of noncentral F (also described by Cohen, 1988, p. 550). The results of this approximation were cross-checked with Bavry's (1991, p. 127) calculation of the noncentral F distribution.

The power of each ANOVA F-test to detect 0.10, 0.20, 0.30, and 0.40 effect size differences is listed in Appendix J. The power of the 2 x 2 x 6 x 6 ANOVA day x test, group x day x test, day x question, group x day x question, test x question, group x test x question, day x test x question, and group x day x test x question F-tests to detect an effect size of 0.20 was at least 0.80 - using a significance criterion of 0.05. The 2 x 2 x 6 x 6 ANOVA test, group x test, question, and group x question F-tests had a power of 0.80 to detect an effect size of 0.30 using a significance criterion of 0.05. The 2 x 2 x 6 x 6 ANOVA had relatively low power to detect group, test, and group x test effect sizes due to the small number of observations in these analyses. The power of reported statistical differences was at least 0.80 at a critical significance level of 0.05 or less. The degrees of freedom used during power calculation were adjusted to compensate for possible violation of sphericity assumptions using the Greenhouse and Geisser epsilon (Є; Geisser & Greenhouse, 1958; Greenhouse & Geisser, 1959; Winer, 1971, p. 523), as suggested by Keppel (1991, pp. 355-356).

Results

Pn1-LnL (Upper Pneumograph Line Length)

Pn1-LnL changed significantly over repeated tests [$F(5, 195) = 3.35$, $Є = .70$, $p = .015$]. Planned comparison results indicated that the average Pn1-LnL measured during test 1 was 51.27 (Average SEM = 10.67) longer [$F(1, 39) = 9.981$, $p = .003$] than the average of those measured during tests 2, 3, 4, 5, and 6. This difference is illustrated in Figure 1-A. The group x question interaction was also significant [$F(5, 195) = 2.84$, $Є = .60$, $p = .041$].

The deceptive and non-deceptive subject responses were analyzed separately to facilitate interpretation of the group x question interaction (Keppel, 1991, pp. 383-384). A significant question effect [$F(5, 98) = 3.59$, $Є = .39$, $p = .038$] was found among the deceptive subject responses, but not among those of the non-deceptive subjects. The results of subsequent comparisons among deceptive subject responses to questions, illustrated in Figure 2-A, indicated that the average response to the question concerning the number 64 was 47.19 (Average SEM = 14.23) shorter [$F(1, 20) = 17.13$, $p = .000$] than the average of the remaining question responses.

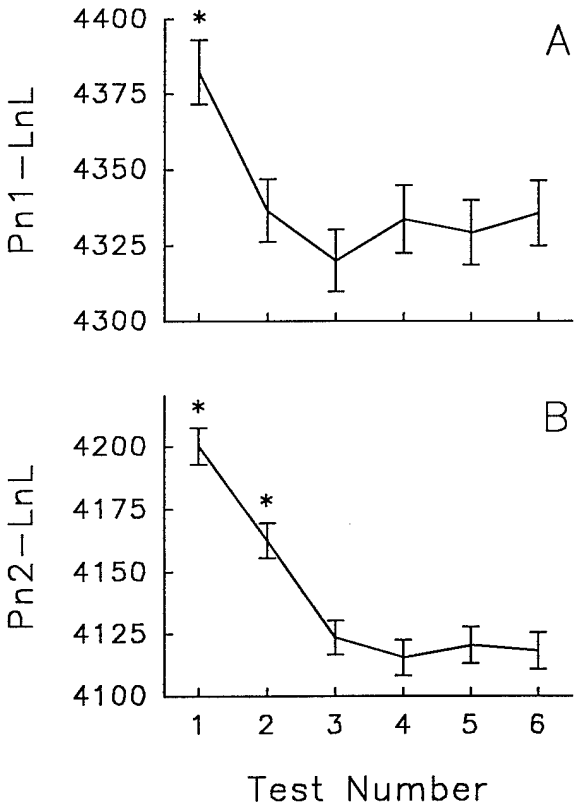


Figure 1. Mean (SEM) response levels for A) Pn1-LnL and B) Pn2-LnL averaged over questions, days, and groups. Values marked with an asterisk (*) are significantly greater than subsequent values.

Pn2-LnL (Lower Pneumograph Line Length)

Pn2-LnL responses measured from the deceptive subjects were an average of 101.76 (Average SEM = 4.03) longer than those measured from non-deceptive subjects [$F(1, 39) = 9.40$, $p = .004$]. Pn2-LnL also changed significantly over repeated tests [$F(5, 193) = 14.89$, $\epsilon = .83$, $p = .000$]. Results of planned comparisons indicated that the average Pn2-LnL measured during test 1 was 80.82 (Average SEM = 7.19) longer [$F(1, 39) = 46.03$, $p = .000$] than the average Pn2-LnL of subsequent tests, and that the average Pn2-LnL measured during test 2 was 72.16 (Average SEM = 7.19) longer [$F(1, 39) = 18.02$, $p = .000$] than the average measured during tests 3, 4, 5, and 6, as illustrated in Figure 1-B. While a significant question effect was found [$F(5, 193) = 3.76$, $\epsilon = .82$, $p = .005$], the planned contrasts were all non-significant. The group x question interaction was also significant [$F(5, 193) = 5.07$, $\epsilon = .82$, $p = .000$].

The deceptive and non-deceptive subject responses were analyzed separately to facilitate interpretation of the group x question interaction. A result of these analyses was that responses were shown to change significantly over repeated tests for both groups. The results of subsequent comparisons among tests showed the same pattern of significant effects as the overall analysis. Responses measured from the deceptive subjects differed significantly during question repetition [$F(5, 97) = 5.52$, $\epsilon = .66$, $p = .000$], while those measured from the non-deceptive subjects did not. Comparison results, illustrated in Figure 2-B, indicate that the deceptive subjects' average Pn2-LnL response to the question concerning the number 64 was 27.96 (Average SEM = 9.82) shorter [$F(1, 19) = 9.05$, $p = .007$] than those in response to the remaining questions. In addition, the deceptive subjects' average Pn2-LnL response to the question concerning the number 65 was 27.43 (Average SEM = 9.82) longer

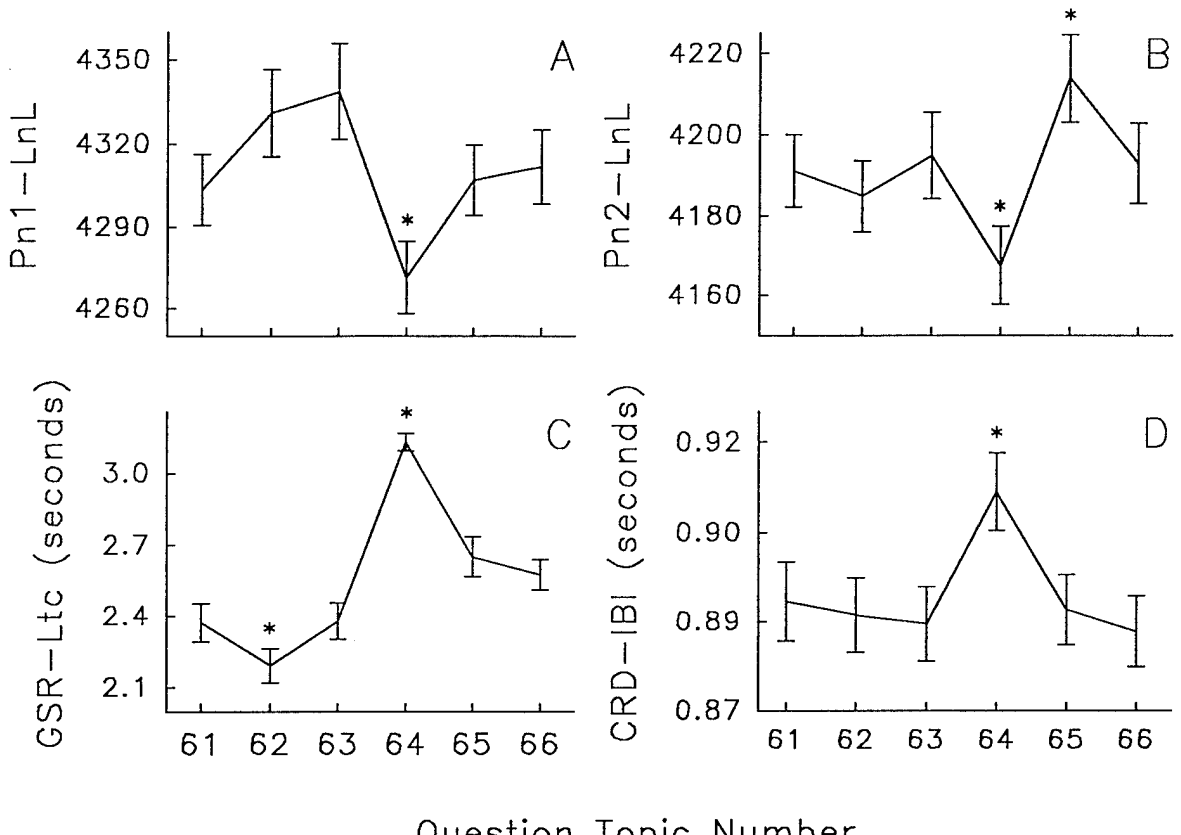


Figure 2. Deceptive subjects' mean (SEM) response levels for A) Pn1-LnL, B) Pn2-LnL, C) GSR-Ltc, and D) CRD-IBI averaged over tests and days. Values marked with an asterisk (*) are significantly greater or less than the average of the remaining values.

[$F(1, 19) = 11.04$, $p = .004$] than the average Pn2-LnL responses to the remaining questions. Responses measured from non-deceptive subjects were also found to differ significantly during question repetition [$F(5, 95) = 3.09$, $\epsilon = .65$, $p = .030$], but no significant differences were found among the subsequent planned comparisons.

GSR-LnL (Galvanic Skin Response Line Length)

A significant group x day x chart interaction [$F(5, 167) = 3.49$, $\epsilon = .86$, $p = .007$] was found among the GSR-LnL measures, but simple effect analysis did not reveal where the differences occurred.

GSR-Amp (Galvanic Skin Response Amplitude)

The average GSR-Amp measured from the deceptive subjects was 10.56 (Average SEM = .877) greater [$F(1, 33) = 10.35$, $p = .002$] than that measured from the non-deceptive subjects. Average GSR-Amp responses also changed significantly [$F(5, 165) = 3.21$, $\epsilon = .85$, $p < .013$] among repeated tests. Planned comparisons, however, failed to reveal any significant

differences. Significant group x question [$F(5, 165) = 13.29$, $\epsilon = .79$, $p = .000$] and group x day x chart [$F(5, 165) = 3.49$, $\epsilon = .84$, $p = .008$] interactions were also found.

Separate analyses of the deceptive and non-deceptive subject GSR-Amp responses were calculated to facilitate interpretation of the group x question and group x day x chart interactions. A significant difference was found among the question responses of the non-deceptive subjects [$F(3, 83) = 9.71$, $\epsilon = .50$, $p = .000$]. Planned comparisons indicated that the average GSR-Amp recorded in response to the question concerning the number 62 was 9.50 (Average SEM = 1.70) greater than the average GSR-Amp response to the remaining questions [$F(1, 16) = 11.34$, $p = .004$]. The average GSR-Amp recorded in response to the question concerning the number 63 was 5.37 (Average SEM = 1.70) less than the average response to the remaining questions [$F(1, 16) = 13.51$, $p = .002$]. Significant differences were also found among the average question [$F(5, 80) = 6.92$, $\epsilon = .74$, $p = .000$] and test [$F(5, 80) = 2.81$, $\epsilon = .74$, $p = .021$] responses of the deceptive subjects. The deceptive subject average GSR-Amp response (Average SEM = 2.39) to the question concerning the number: 62 was 12.73 smaller [$F(1, 16) = 22.25$, $p = .000$] than that to the remaining questions; and, 66 was 10.46 smaller [$F(1, 16) = 16.79$, $p = .000$] than that to the remaining questions. No significant differences were found among the planned comparisons for the tests.

GSR-Ltc (Galvanic Skin Response - Response Latency)

A significant GSR-Ltc measure difference was found among responses to the questions asked during testing [$F(5, 115) = 9.29$, $\epsilon = .84$, $p = .000$]. Comparisons indicate that response latencies to the question concerning the number 63 were 0.23 seconds (Average SEM = .057) shorter [$F(1, 23) = 11.93$, $p = .002$] than those to the remaining questions. Response latencies to the question concerning the number 64 were 0.43 seconds (Average SEM = .057) longer [$F(1, 23) = 49.33$, $p = .000$] than the average of those recorded in response to questions concerning the numbers 61, 62, 63, 65, and 66. The 2 x 2 x 6 x 6 ANOVA also indicated that there was a significant group x question effect [$F(5, 115) = 8.62$, $\epsilon = .84$ $p = .000$].

Data recorded from the deceptive and non-deceptive groups were analyzed separately to assist in interpreting the significant group x question effect. Significant question effects were found for both the non-deceptive [$F(5, 60) = 5.01$, $\epsilon = .65$, $p = .001$] and deceptive [$F(5, 56) = 19.69$, $\epsilon = .76$, $p = .000$] subject responses. No significant differences were found among the question effect planned comparisons for the non-deceptive group. The average deceptive subject GSR-Ltc response (Average SEM = .069) to the question concerning the number 62 was 0.43 seconds shorter [$F(1, 11) = 33.75$, $p = .000$] than the average response to the remaining questions. The average

deceptive subject GSR-Ltc response to the question concerning the number 64 was 0.69 seconds longer [$F(1, 11) = 105.44$, $p < .000$] than the average response to the remaining questions. These differences are illustrated in Figure 2-C.

A significant day x test x question effect [$F(25, 281) = 2.88$, $\epsilon = .35$, $p = .000$] was found among the responses of the deceptive subjects. Separate analyses were calculated for the deceptive subject responses recorded during test days 1 and 2 to assist in interpreting this effect. These analyses indicated significant differences among the GSR-Ltc question responses for both day 1 [$F(5, 56) = 6.07$, $\epsilon = .71$, $p = .001$] and day 2 [$F(5, 56) = 10.20$, $\epsilon = .68$, $p = .000$]. Planned comparisons indicated that the deceptive subject GSR-Ltc responses to the question concerning the number 64, during day 1, were .65 (Average SEM = .10) seconds longer [$F(1, 11) = 41.77$, $p = .000$] than the average response latency to the remaining questions. Comparisons for deceptive responses measured during day 2 (Average SEM = .09) indicate that responses to the question 62 were .38 seconds shorter [$F(1, 11) = 59.36$, $p = .000$] than the average response latency to the remaining questions and that responses to the question concerning the number 64 were .72 seconds longer [$F(1, 11) = 41.37$, $p = .000$] than the average response latency to the remaining questions.

A significant test x question effect was found among the deceptive subjects GSR-Ltc responses during test day 2 [$F(25, 281) = 2.22$, $\epsilon = .32$, $p = .033$]. Each test was analyzed separately to assist in interpreting this difference. No significant differences were found among the question responses recorded during tests 1 and 5. The analyses indicated that there were significant differences among responses recorded to questions during tests 2, 3, 4, and 6. Contrasts indicate that the average GSR-Ltc responses to the question concerning the number 64 were significantly longer ($p < .05$) than the average of those recorded in response to the remaining questions during tests 2, 3, 4, and 6. GSR-Ltc responses to questions concerning the numbers 62 and 66 recorded during test 4 and to questions concerning the number 62 recorded during test 6 were significantly shorter ($p < .05$) than the average of the responses recorded during the remaining questions.

Separate analyses were calculated for the non-deceptive subject responses recorded during test days 1 and 2 to assist in interpreting a significant day x test effect result found during the analysis of non-deceptive subject GSR-Ltc responses [$F(5, 60) = 2.72$, $\epsilon = .68$, $p = .050$]. No significant test, question, or test x question effects were found among the non-deceptive subject GSR-Ltc responses recorded during day 1. Non-deceptive subject responses on day 2 were, however, found to differ significantly among questions [$F(5, 60) = 4.46$, $\epsilon = .623$, $p = .002$]. Planned comparisons indicate that the average GSR-

Ltc response latency to the question concerning the number 63 was .37 (Average SEM = .12) seconds shorter than the average response latency to the remaining questions [$F(1, 12) = 13.71$, $p = .003$].

CRD-IBI (Cardio Channel Average Inter-beat-interval)

A significant CRD-IBI measure difference was found among responses to the questions asked during testing [$F(5, 197) = 4.27$, $\epsilon = .53$, $p = .009$]. Comparisons indicate that the average CRD-IBI measured in response to the question concerning the number 64 was 0.011 (Average SEM = .005) seconds longer [$F(1, 39) = 14.80$, $p = .000$] than those to the remaining questions. The 2 x 2 x 6 x 6 analysis also indicated significant group x question [$F(5, 197) = 3.41$, $\epsilon = .53$, $p = .025$], group x day x test [$F(5, 197) = 3.06$, $\epsilon = .83$, $p = .017$], and group x test x question interactions [$F(25, 987) = 1.93$, $\epsilon = .45$, $p = .03$].

Separate analyses of the data recorded from the deceptive and non-deceptive subjects were calculated to facilitate interpretation of the significant interaction effects. The analysis indicated no significant differences among the non-deceptive subject responses as a function of the independent variables manipulated. A significant question effect was, however, found among the deceptive subject responses [$F(5, 99) = 5.84$, $\epsilon = .54$, $p = .002$]. Planned comparisons indicated that the deceptive subjects average CRD-IBI response to the question concerning the number 64 was .021 (Average SEM = .0061) seconds longer than the average response CRD-IBI to the remaining questions, as illustrated in Figure 2-D.

Discussion

Interpretation of these results suggests that during repeated administration of PDD tests: there is a consistent change in average Pn1-LnL and Pn2-LnL; differential Pn1-LnL, Pn2-LnL, and CRD-IBI reactivity during a PDD test does not change during repeated tests or days; and, average physiological reactivity of deceptive subjects changes during deception while that of non-deceptive subjects does not. When interpreting these results it is important to remember that the power of each significant statistical effect was 0.80 or greater and that the power of the non-significant statistical tests to detect an effect of size 0.30 at the 0.05 significance level was also 0.80 or greater (with exceptions noted above). The power analysis provides the probability (0.80 or greater) that the null hypothesis is correctly rejected when a significant effect was observed, as well as the probability (0.80 or greater) that an effect size of 0.30 would have been correctly detected.

Perhaps the most interesting result of this research is not the significant results which were obtained, but those that were not. All day x test, day x question, test x question, and day x

test x question interactions were non-significant. This suggests that the pattern and/or variability of measured physiologic responses to the questions asked during each PDD test did not change significantly over repeated administration of the tests, nor did the response pattern change significantly between days 1 and 2 - with the exception of GSR-Ltc responses. This result is interpreted as supporting those of Ellson et al. (1952), Lieblich et al. (1974), Grimsley and Yankee (1986), and Yankee (1993) that there were no statistically significant differences in the detection of veracity with repeated testing. While veracity detection rates were not determined, the conclusion that differential responding does not change with question series repetition supports the proposal that decision accuracy does not decrease with repeated testing (Grimsley & Yankee, 1986; Iacono, Boisvenu, & Fleming, 1984; Leiblich et al., 1974).

The results of some investigations into the effect of repeated question series administration on skin resistance and/or conductance responsivity do not support those of this study (Balloun & Holmes, 1979; Ben-Shakhar & Lieblich, 1982; Elaad & Ben-Shakhar, 1989; Iacono et al., 1984) while those of others do (Furedy, Ben-Shakhar, 1991; Furedy, Gigliotti, & Ben-Shakhar, 1994). This is a difficult issue to resolve due to methodological differences in the: response requirements; question repetition patterns and procedures; and, data reduction, evaluation, and analysis techniques. It is also possible that the response strengths measured during this study decreased with repetition, but the decrease was too small to be statistically detected. It is likely, however, that such small changes would be of little interest. Further research should be conducted to address these issues.

Average Pn1-LnL and Pn2-LnL response levels measured during the first test, averaged over groups, days, and questions, were found to be significantly greater than the average of the subsequent tests, as illustrated in Figure 1. No statistically significant difference was found between Pn1-LnL measures recorded during tests 2 through 5 and the average of subsequent tests. The average Pn2-LnL measure recorded during test 2 was significantly greater than the average recorded during tests 3 through 6, but measures recorded during tests 3 through 5 were no different from those recorded during subsequent tests. A similar shift in skin conductance following repeated testing has been reported by Iacono et al. (1984). The decrease in average response levels observed during the initial stages of repeated testing, in the absence of within test response attenuation, may be a variation of the phenomenon of differential autonomic responsivity, proposed by Ben-Shakhar and Lieblich (1982).

Results of the data analyses indicate that there were no statistically significant main or interaction effects related to

the questions asked among the average non-deceptive subject Pn1-LnL, Pn2-LnL, and CRD-IBI responses. The average deceptive subjects' deceptive responses were shorter in Pn1-LnL and Pn2-LnL, longer in GSR-Ltc, and longer in CRD-IBI than the average of their non-deceptive responses. These results confirm that, on the average, a pattern of differential responding occurs during deception that does not occur when deception is not present. While pneumo line lengths and heart rate are not normally evaluated when scoring PDD examinations, perhaps polygraphs used for PDD should be modified to display this information.

While significant differences were found among the deceptive subjects' GSR-Amp responses to the questions asked, the deceptive response was not significantly different from the average non-deceptive response. This is surprising when one considers results of studies reporting high veracity detection accuracy rates based exclusively on electrodermal activity scores (Iacono, Cerri, Patrick, & Fleming, 1992; Kugelmass & Lieblich, 1966; Podlesny & Raskin, 1978; Thackary & Orne, 1968). However, close examination of these reports suggests that differences in methodology and evaluation techniques could account for the differences between the current results and earlier reports. While a field polygraph was used in the current study, the operator sensitivity adjustments were bypassed. Skin resistance changes were amplified by a fixed-gain linear amplifier adjusted to remain within the range limits of an analog-to-digital converter, which did not compensate for changes in tonic skin resistance, possibly contributing to the failure to find significant differences among GSR-Amp measures during deception, in this study.

It should, however, be noted that 9% and 27% of the subjects were dropped from the GSR-Amp and GSR-Ltc analyses, respectively, due to insufficient data caused, primarily, by failure to obtain quantifiable subject responses. The percentages of missing Pn1-LnL, Pn2-LnL, and CRD-IBI data, which were collected simultaneously with the GSR data, were not sufficiently large to necessitate removal of subjects from the analyses. This observation is interpreted as suggesting that the exclusive or disproportionately high reliance on GSR response scores when interpreting the results of PDD examinations may lead to excessive errors. This suggestion is not new, but simply reinforces the statement presented to the Committee on Government Operations over 20 years ago that "most examiners agree that the galvanic skin response is the least accurate, and should be ignored when a conflict (among the three channels) occurs" (Committee on Government Operations, 1974, p. 24).

In summary, three conclusions are derived from the results of this research. First, a consistent change was observed in

average Pn1-LnL and Pn2-LnL responses, but not the GSR-Amp, GSR-LnL, GSR-Ltc, and CRC-IBI responses as the test was repeated. This pattern did not change significantly between test days one and two. Second, the average physiological response variability measured during a PDD test did not change over repeated tests. Finally, the Pn1-LnL, Pn2-LnL, GSR-Ltc, and CRD-IBI responses of deceptive subjects, averaged over repeated test administrations, changed during the deceptive response, relative to non-deceptive responses. No such systematic changes were found among the responses of the non-deceptive subjects. These data are interpreted as suggesting that decision accuracy will not decrease significantly during repeated (up to six) administrations of the question series during a PDD examination. This conclusion is supported by other reports (Grimsley & Yankee, 1986; Iacono et al., 1984; Leiblich et al., 1974). It is suggested that changes in heart rate inter-beat-interval, measured using an occlusive cuff as described, and pneumo line length are reliable response measures which may be accurately interpreted as indicating deception.

References

Abrams, S. (1989). The complete polygraph handbook. Lexington, MA: Lexington Books.

Balloun, K. D., & Holmes, D. S. (1979). Effects of repeated examinations on the ability to detect guilt with a polygraphic examination: a laboratory experiment with a real crime. Journal of Applied Psychology, 64, 316-322.

Barland, G. H., & Raskin, D. C. (1975). An evaluation of field techniques in detection of deception. Psychophysiology, 12, 321-330.

Bavry, J. L. (1991). Stat-Power statistical design analysis system. Chicago, IL: Scientific Software, Inc.

Ben-Shakhar, G., & Lieblich, I. (1982). The dichotomization theory for differential autonomic responsivity reconsidered. Psychophysiology, 19, 277-281.

Bersh, P. J. (1969). A validation study of polygraph examiner judgments. Journal of Applied Psychology, 53, 399-403.

Campbell, R. J. (1989). Psychiatric Dictionary (6th ed). New York: Oxford University Press.

Cestaro, V. L., & Dollins, A. B. (1994). An analysis of voice responses for the detection of deception (Report No. DoDPI94-R-0001). Ft. McClellan, AL: Department of Defense Polygraph Institute.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Committee on Government Operations (1974). The use of polygraphs and similar devices by federal agencies (Hearings before a Subcommittee of the Committee on Government Operations, House of Representatives, 93rd Congress, 2nd Session). Washington, DC: U.S. Government Printing Office.

Elaad, E., & Ben-Shakhar, G. (1989). Effects of motivation and verbal response type on psychophysiological detection of information. Psychophysiology, 26, 442-451.

Ellson, D. G., Davis, R. C., Saltzman, I. J., & Burke, C. J. (1952). Accuracy of detection and the effect of repetition. A report of research on detection of deception (Tech. Rep. Contract No. N60NR-18011). Bloomington, IN: Indiana University, Department of Psychology.

Furedy, J. J. (1986). Lie detection as a psychophysiological differentiation: some fine lines. In M. G. H. Coles, E. Donchin, & S. W. Porges (Eds.), Psychophysiology: Systems, Processes, and Applications (pp. 683-701). New York: Guilford Press.

Furedy, J. J., & Ben-Shakhar, G. (1991). The roles of deception, intention to deceive, and motivation to avoid detection in the psychophysiological detection of guilty knowledge. Psychophysiology, 28, 163-171.

Furedy, J. J., Gigliotti, F., & Ben-Shakhar, G. (1994). Electrodermal differentiation in deception: the effect of choice versus no choice of deceptive items. International Journal of Psychophysiology, 18, 13-22.

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. Annals of Mathematical Statistics, 29, 885-891.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. Psychometrika, 24, 95-112.

Grimsley, D. L., & Yankee, W. J. (1986). The effect of multiple retests on examiner decisions in applicant screening polygraph examinations (National Security Agency Contract No. MDA 904-85-C-A962). Charlotte, NC: A. Madley Corporation.

Horvath, F., & Reid, J. (1971). The reliability of polygraph examiner diagnosis of truth and deception. Journal of Criminal Law, Criminology, and Police Science, 62(2), 276-281.

Horvath, F. (1977). The effect of selected variables on interpretation of polygraph records. Journal of Applied Psychology, 62, 127-136.

Hunter, F. L., & Ash, P. (1973). The accuracy and consistency of polygraph examiners' diagnoses. Journal of Police Science and Administration, 1, 370-375.

Iacono, W. G., Boisvenu, G. A., & Fleming, J. A. (1984). Effects of diazepam and methylphenidate on the electrodermal detection of guilty knowledge. Journal of Applied Psychology, 69, 289-299.

Iacono, W. G., Cerri, A. M., Patrick, C. J., & Fleming, J. A. E. (1992). Use of antianxiety drugs as countermeasures in the detection of guilty knowledge. Journal of Applied Psychology, 77, 60-64.

Keppel, G. (1991). Design and analysis, a researcher's handbook (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Koele, P. (1982). Calculating power in analysis of variance. Psychological Bulletin, 92, 513-516.

Kugelmass, S., & Lieblich, I. (1966). The effects of realistic stress and procedural interference in experimental lie detections. Journal of Applied Psychology, 50, 211-216.

Laubscher, N. F. (1960). Normalizing the noncentral t and F distributions. Annals of Mathematical Statistics, 31, 1105-1112.

Lieblich, I., Naftali, G., Shumueli, J., & Kugelmass, S. (1974). Efficiency of GSR detection of information with repeated presentation of series of stimuli in two motivational states. Journal of Applied Psychology, 59, 113-115.

Lykken, D. T. (1960). The validity of the guilty knowledge technique: the effects of faking. Journal of Applied Psychology, 44, 258-262.

Lykken, D. T. (1981). A Tremor in the Blood: Uses and Abuses of the Lie Detector. New York: McGraw-Hill.

Office of Technology Assessment (1983). Scientific Validity of Polygraph Testing: A Research Review and Evaluation - A Technical Memorandum (Report No. OTA-TM-H-15, November). Washington, DC: Office of Technology Assessment.

Podlesny, J., & Raskin, D. (1978). Effectiveness of techniques and physiological measures in detection of deception. Psychophysiology, 15, 344-359.

Reid, J. E. (1947). A revised questioning technique in lie detection tests. Journal of Criminal Law & Criminology, 37, 542-547.

Slowik, S. M., & Buckley, J. P. (1975). Relative accuracy of polygraph examiner diagnosis of respiration, blood pressure, and GSR recordings. Journal of Police Science and Administration, 3, 305-309.

Thackray, R., & Orne, M. (1968). A comparison of physiological indices in detection of deception. Psychophysiology, 4, 329-339.

Timm, H. (1979). The effect of placebos and feedback on the detection of deception (U.S. Department of Justice Contract No. 78-NI-AX-0028). East Lansing, MI: Michigan State University.

Timm, H. (1982a). Analyzing deception from respiration patterns. Journal of Police Science and Administration, 10(1), 47-51.

Timm, H. (1982b). Effect of altered outcome expectancies stemming from placebo and feedback treatments on the validity of the guilty knowledge technique. Journal of Applied Psychology, 67(4), 391-400.

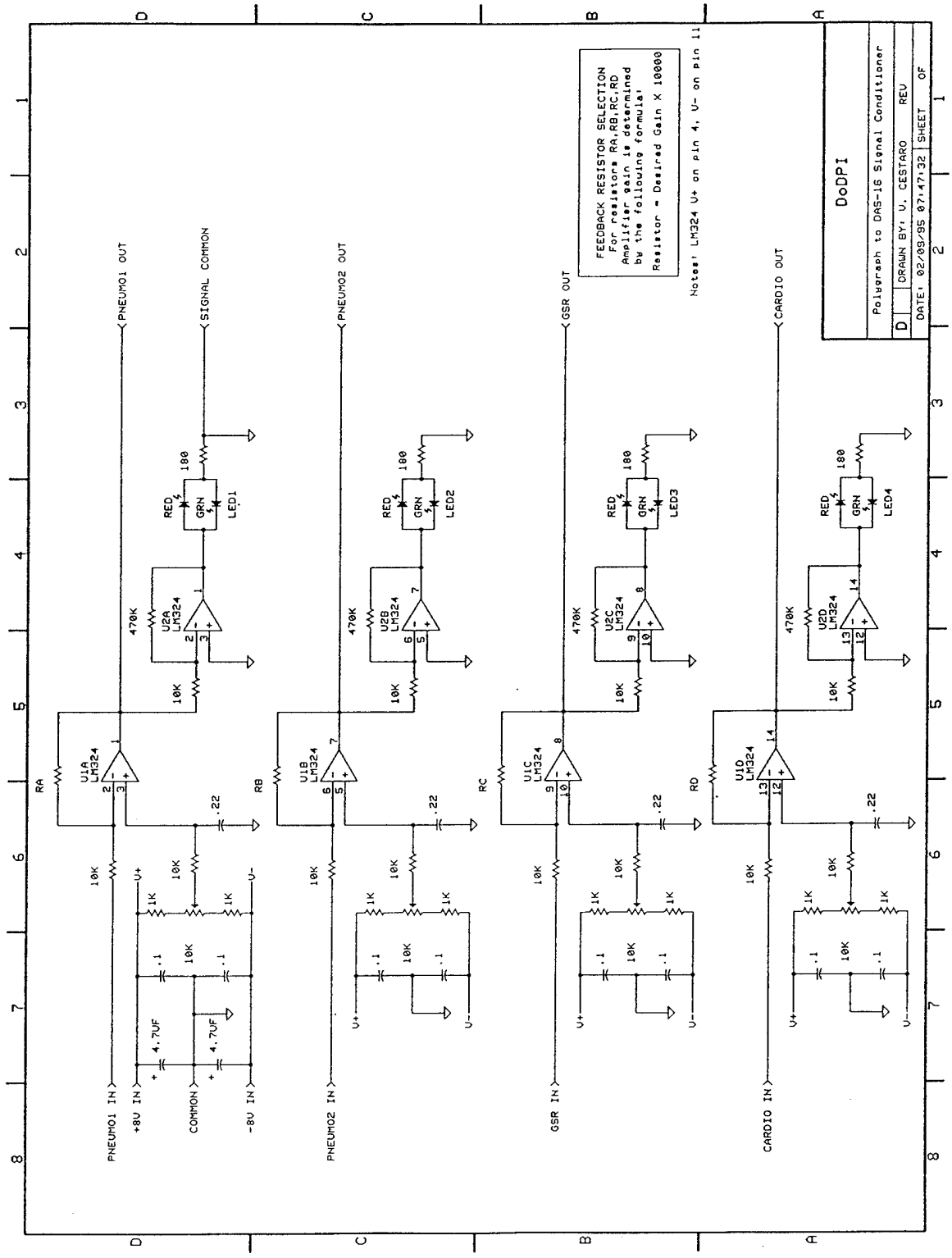
Williams, R. H., & Zimmerman, D. W. (1989). Statistical power analysis and reliability of measurement. Journal of General Psychology, 116, 359-369.

Winer, B. J. (1971). Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill.

Yankee, W. J. (1993). Test and re-test accuracy with a psychophysiological detection of deception test. Department of Defense Polygraph Institute, Ft. McClellan, AL. Adopted from Yankee, W. J. & Grimsly, D. L. (1986) The effect of a prior polygraph test on a subsequent polygraph test (NSA contract No. MDA 904-84-C-4249). Charlotte, NC: University of North Carolina.

Appendix A

Schematic Representation of Constant Gain Linear Amplifier Used to Bypass Polygraph Operator Sensitivity and Centering Adjustments



Appendix B

Description of Research

WELCOME: Welcome to the Department of Defense Polygraph Institute. This may be the first time you have been to the Institute so we would like to provide you with some information concerning your visit today. PLEASE REMEMBER that your participation is entirely voluntary - you are free to leave at any time. If you have any questions, please feel free to ask the individuals assisting you.

Research Title: Efficacy of Repeated Psychophysiological Detection of Deception Testing (June 9, 1993)

Principal Investigator: Dr. Andrew B. Dollins, DoDPI Research Psychologist

BACKGROUND / SIGNIFICANCE: The Psychophysiological Detection of Deception (PDD) is a process believed to determine whether an individual is responding truthfully to a series of questions. PDD is commonly called "lie detection" or "polygraph" test. The process is based on the assumption that an individual who is deceptive (i.e., lying) has a greater response in some body systems than a person who is not. While this is generally true, it is not known whether an individual always responds in the same way when being deceptive. This project is designed to test the consistency of responses when an individual is lying and is telling the truth.

YOU SHOULD NOT PARTICIPATE IN THIS STUDY IF YOU:

1) Previously participated in a PDD examination.
2) Are taking prescription medication.
3) Have a history of dizziness or fainting spells.
4) Have been diagnosed with a heart condition.
5) Have been diagnosed with high blood pressure.
6) Have been diagnosed with a respiratory ailment, especially asthma or emphysema.
7) Currently suffer from an acute health problem such as a cold, active allergy problem, hemorrhoidal problem.

PROCEDURES: During this project you will be asked to participate in two research sessions lasting approximately four hours each. These two sessions will be separated by five to ten days as scheduling permits. During each session you will be asked to complete a puzzle and, possibly, to lie about the puzzle during a PDD examination. Some people will be asked to lie about the puzzle they completed and some will not be asked questions about the puzzle. If you are asked questions about the puzzle you completed, YOUR TASK IS TO LIE SUCCESSFULLY, to the PDD examiner concerning the puzzle.

Participation in the PDD process is relatively simple. The examiner will ask several questions concerning your age, health, and normal daily activities. He will then briefly explain the theory of the Psychophysiological Detection of Deception and review the questions he will ask during the examination with you. With your permission, the examiner will then attach sensors to your body. Two small flat metal sensors will be attached to the first and third

fingers of one hand. Expandable tubes will be put around your upper and lower chest. A blood pressure cuff will be wrapped around your arm. You will be asked to sit still for several minutes while the examiner asks the questions he reviewed earlier. The examiner may ask the same questions several times during the examination. When the session is complete, we will make an appointment for your second session and you will be escorted out of the building. The second session will be like the first except this briefing will not be repeated.

DISCOMFORTS: Some people find it difficult to sit still for several minutes at a time during the PDD test while physiological reactions are recorded. Part of the PDD process requires the wearing of an inflated blood pressure cuff, which some people find moderately uncomfortable. The examiner is sensitive to this discomfort and will attempt to make the process as brief as possible. The actual tests last approximately five minutes each. You will be asked to participate in as many as nine tests during each examination day. The total length of time that you will actually be participating in a polygraph examination is 45 minutes to two hours, however, you may be at DoDPI for three or four hours.

VIDEOTAPING: All examinations conducted during this project will be videotaped using wall and ceiling mounted video cameras and commercial videotape recorders. The tapes collected will be maintained until the operational and data analysis portions of the project are complete. At that time the video tapes will be erased and made available for re-use by the research and instruction divisions.

RISKS: There are no known risks involved in this study.

CONFIDENTIALITY OF RECORDS: You will not be asked any personal questions by the examiner, except medically related information necessary for this study. Neither your identity nor any information you reveal during this project will be released to anyone not directly involved in the research. Members of the U.S. Army Surgeon General's Human Subjects Research Review Board may inspect the research records in their capacity as reviewing officials.

YOUR RIGHTS: You have the right to ask any questions about any aspect of your participation in the study. If any problems arise at any time in conjunction with your involvement in the study, or if you have been injured in any way as a result of the study, the person to contact is the Chief of Research, Department of Defense Polygraph Institute. In the event that you do have questions or any of the above has occurred please contact Dr. William Yankee at (205) 848-3803. Should any question arise concerning study-related injury, you may contact the Director of the Noble Army Community Hospital, Fort McClellan, Alabama, 36205, telephone number (205) 848-2200.

VOLUNTARY PARTICIPATION: Your participation in this study is completely voluntary. **If you would prefer not to participate, do not volunteer for it!** Even if you decide to participate in the study, you may discontinue at any time without penalty or loss of benefits to which you are entitled. Should you decide not to participate, please inform someone on the staff at the Department of Defense Polygraph Institute, or if it occurs during the

polygraph examination itself, inform the examiner and you will be released without censure.

ADDITIONAL COMMENTS: Regardless of whether you are required to lie during the PDD examination, it is very important that you do not tell the examiner whether you are being truthful or not. Examiners should not ask and if they do, please tell another staff member. It is also VERY IMPORTANT that you do not discuss your experiences in the PDD examination with your fellow research participants. If either of the above occurs, you will be withdrawn from the study without further benefit.

Appendix C

Volunteer Aqreement Affidavit

This form is affected by the Privacy Act of 1974.

1. **AUTHORITY:** 10 USC 3012, 44 USC 3101 and 10 USC 1071-1087.

2. **PRINCIPLE PURPOSE:** To document voluntary participation in a DoD Polygraph Institute Research Program.

3) **ROUTINE USES:** Your name will be used for identifying and locating research documents and will be available only to individuals associated with the research project.

4. **MANDATORY OR VOLUNTARY DISCLOSURE:** Your signature is necessary if you want to be included in this research. If you do not sign, you will not be able to participate in this study and you will not be paid.

PERSONAL STATEMENT

I, _____, being at least 19 years old, do hereby volunteer to participate in a research study titled "Efficacy of Repeated Psychophysiological Detection of Deception Testing" being conducted at the Department of Defense Polygraph Institute, under the direction of Andrew B. Dollins, Ph.D.

1._____ I understand that I am participating in a research study to examine several measures and techniques, some of which are currently employed in criminal and/or security screening situations where the Psychophysiological Detection of Deception (PDD) is used. PDD is commonly called a 'polygraph test' or 'lie detector'.

2. To the best of my knowledge,

A._____ I am not taking any prescription medication.

B._____ I have no history of dizziness or fainting spells.

C._____ I have not been diagnosed as having, nor do I believe that I may have any of the following:

1) Heart condition.

2) High blood pressure.

3) Any respiratory ailment, especially asthma or emphysema.

D._____ I do not now have any acute health problems such as a cold, an active allergy problem, and an active hemorrhoidal problem.

3._____ I am aware that I will be spending approximately four (4) hours at the Department of Defense Polygraph Institute (DoDPI) on two occasions, and that I may be asked to conceal information concerning my activities at DoDPI from a trained Forensic Psychophysiologist.

4._____ I understand that as a part of this study I will be participating in a PDD examination during which I will be asked to sit still for several minutes at a time while physiological measurements are recorded from my body.

5._____ I understand that there are no known dangers or risks associated with my participation in this study.

6._____ I understand that I will be required to wear an inflated blood pressure cuff, which some people find moderately uncomfortable, during the PDD examination.

7._____ I understand that I will be videotaped during the PDD examination and that the videotape will be maintained until data analyses are complete.

8._____ I understand that I will receive no reward or benefit of any kind as a result of my participation in this study.

9._____ I understand that I may terminate my involvement in this study at any time and for any reason, without censure.

10._____ I understand that my participation in this project will be terminated if I discuss the details of my participation with anyone except project supervisory personnel. NOTE: Discussion of details with other participants would invalidate the data collection.

11._____ I understand that I should contact the principal investigator, Dr. Andrew Dollins, and / or the DoD Polygraph Institute Director, Dr. William Yankee [Telephone number: (205) 848-3803] if I have any concerns or complaints regarding this study.

12._____ I understand that any questions concerning my rights relating to study-related injury should be directed to Colonel Weisser, MD, Director of the Noble Army Community Hospital, Fort McClellan, Alabama, 36205, telephone number (205) 848-2200.

13._____ I have been given a thorough explanation of the nature, purpose, methods, and duration of my participation in this investigation. I have been given the opportunity to ask any questions I have concerning the investigation and all questions have been answered to my full satisfaction.

_____	_____
Participant Signature	Witness Signature
_____	_____
Printed Name	Printed Name
_____	_____
Date	Date

Appendix D

Biographical/Medical Form

Participant number:_____ Date of completion:_____

Please carefully complete all of the blanks below:

Name (Please Print):_____ Gender:( )M ( )F

Occupation:_____ Age:_____

Hours of sleep last night:_____ Previous PDD Examination: ( )Yes ( )No

Have you ingested alcohol, nicotine, or caffeine (including coffee, tea, soft-drinks, and chocolate) within the last 24 hours? ( )Yes ( )No

If so, what and when?_____

_____

How would you describe your present health and physical well being?
( )Excellent ( )Good ( )Fair ( )Poor

Are you presently under a physician's care and are you taking any medication? ( )Yes ( )No

If so, for what condition?_____

Please identify the type, dosage, and last time any medication was taken:

_____

Are you experiencing any pain or discomfort today?
( )None ( )Mild ( )Moderate ( )Severe

Reason for any pain or discomfort today:_____

_____

Appendix E

Example of Anagram Task Completed by Deceptive Subjects

ANAGRAM TASK

Please locate and circle six sequences of
the number which is repeated five times below.
(See example for the number 22 on the right.)

Name:_____

Subject #:_____ Date:_____

Score:_____ (EoRPDD 07/12/93 6A)

20	21	22	21	29	21	24	22
21	28	25	22	29	21	21	29
25	22	22	22	22	22	28	24
27	25	28	27	25	22	25	24
24	28	22	22	28	25	22	20
24	21	22	24	20	23	26	25
21	22	22	23	22	29	21	21
22	23	22	20	26	28	27	29
24	22	22	28	24	22	26	23

66	62	64	61	61	63	65	64	67	66	66	66	61	64	63	64	65	66	67	62
68	69	63	66	67	61	65	68	68	67	68	68	65	65	65	66	68	63	68	68
68	62	69	62	65	66	64	64	64	64	64	68	69	66	66	66	61	62	67	66
61	64	63	61	63	66	68	69	69	64	69	67	66	66	63	63	65	60	62	65
67	66	67	62	65	64	61	65	61	66	62	62	68	60	66	64	67	62	65	66
68	60	68	69	68	65	63	60	63	69	65	68	67	67	65	64	67	68	66	65
64	69	65	62	60	62	60	65	62	69	68	62	67	61	61	64	67	68	62	63
67	69	65	64	63	69	65	64	62	61	65	61	64	67	66	64	69	65	62	67
65	60	65	61	68	68	60	68	65	66	62	68	61	69	68	64	65	66	61	63
65	68	65	63	64	61	65	62	64	65	62	63	65	67	63	67	63	62	69	63
65	66	64	63	66	64	67	65	64	64	60	60	68	66	64	68	66	62	63	67
67	61	65	60	65	61	61	63	63	67	64	62	61	63	68	61	67	64	67	60
67	68	67	69	64	68	68	61	63	66	64	64	63	67	66	60	66	69	63	61
61	68	66	61	69	69	61	67	69	62	68	67	64	61	64	62	66	66	61	63
62	60	67	61	63	61	68	65	64	63	69	64	63	63	63	65	60	65	64	65
68	68	61	64	63	68	64	62	62	67	67	68	62	63	65	67	60	66	64	63
69	61	62	61	68	61	66	64	65	67	64	60	63	68	68	68	64	63	64	65
62	67	62	61	61	62	67	66	61	65	65	65	62	62	65	69	61	62	64	66
64	62	62	63	67	62	63	67	63	63	60	61	60	63	63	66	60	63	64	62
63	64	64	62	67	66	61	61	63	66	66	66	64	66	64	63	68	67	67	68
68	66	67	64	64	64	64	64	61	61	62	63	66	61	65	65	62	65	64	61
65	65	63	62	69	69	62	68	62	69	68	66	67	65	69	69	61	65	62	63
64	62	65	63	69	67	65	64	67	66	65	65	63	63	65	62	61	68	67	67
64	64	63	67	65	69	64	61	60	68	68	68	62	67	62	65	67	66	66	60
66	62	61	63	62	66	65	62	60	60	67	65	65	60	65	64	63	69	65	67
65	69	67	60	62	67	61	64	63	68	61	65	65	66	66	67	68	60	67	64
64	66	61	66	63	63	64	68	61	68	61	62	61	66	62	64	68	61	61	68
63	69	61	67	63	64	67	67	62	67	67	64	63	69	64	64	68	67	61	61
60	62	62	65	64	68	64	67	61	68	61	67	62	64	63	61	62	62	69	65
60	65	69	64	63	66	61	61	65	64	61	61	64	63	62	61	69	67	67	61

(Form Truncated)

Appendix F

Example of Anagram Task Completed by Non-deceptive Subjects

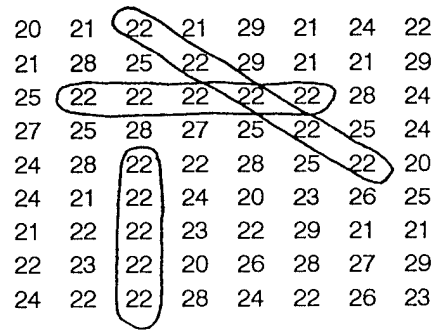
ANAGRAM TASK

Please locate and circle six sequences of
the number which is repeated five times below.
(See example for the number 22 on the right.)

Name:_____

Subject #:_____ Date:_____

Score:_____ (EoRPDD 07/12/93 8A)



86	82	84	81	81	83	85	84	87	86	86	86	81	84	83	84	85	86	87	82
88	89	83	86	87	81	85	88	88	87	88	88	85	85	85	86	88	83	88	88
88	82	89	82	85	86	84	84	84	84	84	88	89	86	86	86	81	82	87	86
81	84	83	81	83	86	88	89	89	84	89	87	86	86	83	83	85	80	82	85
87	86	87	82	85	84	81	85	81	86	82	82	88	80	86	84	87	82	85	86
88	80	88	89	88	85	83	80	83	89	85	88	87	87	85	84	87	88	86	85
84	89	85	82	80	82	80	85	82	89	88	82	87	81	81	84	87	88	82	83
87	89	85	84	83	89	85	84	82	81	85	81	84	87	86	84	89	85	82	87
85	80	85	81	88	88	80	88	85	86	82	88	81	89	88	84	85	86	81	83
85	88	85	83	84	81	85	82	84	85	82	83	85	87	83	87	83	82	89	83
85	86	84	83	86	84	87	85	84	84	80	80	88	86	84	88	86	82	83	87
87	81	85	80	85	81	81	83	83	87	84	82	81	83	88	81	87	84	87	80
87	88	87	89	84	88	88	81	83	86	84	84	83	87	86	80	86	89	83	81
81	88	86	81	89	89	81	87	89	82	88	87	84	81	84	82	86	86	81	83
82	80	87	81	83	81	88	85	84	83	89	84	83	83	83	85	80	85	84	85
88	88	81	84	83	88	84	82	82	87	87	88	82	83	85	87	80	86	84	83
89	81	82	81	88	81	86	84	85	87	84	80	83	88	88	88	84	83	84	85
82	87	82	81	81	82	87	86	81	85	85	85	82	82	85	89	81	82	84	86
84	82	82	83	87	82	83	87	83	83	80	81	80	83	83	86	80	83	84	82
83	84	84	82	87	86	81	81	83	86	86	86	84	86	84	83	88	87	87	88
88	86	87	84	84	84	84	84	81	81	82	83	86	81	85	85	82	85	84	81
85	85	83	82	89	89	82	88	82	89	88	86	87	85	89	89	81	85	82	83
84	82	85	83	89	87	85	84	87	86	85	85	83	83	85	82	81	88	87	87
84	84	83	87	85	89	84	81	80	88	88	88	82	87	82	85	87	86	86	80
86	82	81	83	82	86	85	82	80	80	87	85	85	80	85	84	83	89	85	87
85	89	87	80	82	87	81	84	83	88	81	85	85	86	86	87	88	80	87	84
84	86	81	86	83	83	84	88	81	88	81	82	81	86	82	84	88	81	81	88
83	89	81	87	83	84	87	87	82	87	87	84	83	89	84	84	88	87	81	81
80	82	82	85	84	88	84	87	81	88	81	87	82	84	83	81	82	82	89	85
80	85	89	84	83	86	81	81	85	84	81	81	84	83	82	81	89	87	87	81

(Form Truncated)

Appendix G

Outline of Examiner's Explanation of Sensors, Procedures, and PDD Theory

Good morning (afternoon), my name is (insert name here) and I will be conducting the polygraph examination today. I am an instructor at the Polygraph Institute and like you I have been detailed to assist Dr. Dollins in this very important research project. You and I know that this project is very important otherwise the Army would not have provided us to participate.

Before we begin conducting any examinations I will explain everything that will be attached to you for this examination and we will have discussed a little bit about your background and one of the theories of psychophysiological detection of deception. Let me assure you that nothing will be said or done here that will in any way hurt or injure you. Do you have any questions before we proceed?

Now, I would like to review the interview work sheet.
[Review Pre-Test Questionnaire - Appendix D]

One of the theories concerning the psychophysiological detection of deception or the ability of a trained forensic psychophysiologist (polygraph examiner) to diagnose deception is that of Fight or Flight which you may be familiar with from sports and your training in the military. This phenomenon is theorized to be what allows us to survive in dangerous or stressful situations. When the mind recognizes that we are in danger we enter into Fight or Flight and epinephrine is released into the blood stream. This drug effects different organs of the body in different ways. In the case of the cardiovascular system this drug causes the activity of the heart to increase along with a marked increase in the pulse, blood pressure, and other cardiac activity.

In the case of the heart the increases are to provide more oxygen and nutrients to the large muscles of the legs and arms so we can run away from the problem or fight our way out of the problem. Additionally this provides more oxygen to the brain so we can think our way out of the problem. The epinephrine additionally effects our lungs by causing them to increase activity to better place oxygen in the blood stream and to remove carbon dioxide from the system.

The body experiences numerous other physiological changes to include changes in the sweat gland activity and the electrodermal activity at the skin. Normally these reactions are associated with fear. These reactions are what allows us to survive in stressful situations such as combat, parachuting, and other duties.

[The Examinee is then asked to provide an example of when they might have experienced this phenomenon. Common examples were as follows: 1st traffic citation; combat in South West Asia; traffic accidents; and, training mishaps.]

Well, I can tell by your example that you are familiar with these reactions. The same type of reactions occur when we are practicing deception because there is a fear of being caught in an untruthful statement or being punished for the untruth. Have you ever experienced these reactions?

Once we have told the deception another drug is released into the blood stream which brings the body back to normal. This drug is called nor-epinephrine. This same drug aids in our recovery from dangerous situations.

With the sensitive apparatus associated with a polygraph instrument a trained polygraph examiner can diagnose when an individual has been less than truthful when answering questions while attached to the instrument. The actual attachments that will be placed on your body are the standard hospital blood pressure cuff, to monitor your cardiac activity. Two small metal plates which will be attached to you finger tips to monitor your sweat gland activity, and two convoluted tubes which will be placed around your torso to monitor you respiratory activity. None of these attachments will cause you any pain or discomfort. Also a microphone will be placed around your neck to make an accurate recording of your verbal responses to the questions on today's test.

The examinee is then presented with the prerecorded questions for this examination.

Appendix H

Participant Debriefing Statement I

Now that you have completed your first examination, it is the desire of the entire project staff to take this opportunity to sincerely thank you for your help. Your work here may be more important than you realize.

If you participated in deceiving the PDD examiner, you are assured by the staff of this Institute, that you in no way violated any rule or law. The deception was required for investigational purposes only.

Regardless of the role you played, it is our hope that you were made to feel as comfortable as possible throughout the study. If you do have concerns or questions regarding your participation, please make them known to the principal investigator, Dr. Andrew Dollins, and / or the DoD Polygraph Institute Director, Dr. William Yankee [Telephone number: (205) 848-3803].

Finally, it is VERY IMPORTANT that you DO NOT discuss the details of this study with anyone else. One of your friends, or a friend of a friend, may decide to participate in this or a similar study someday. If they know the details of the investigation process, they could be disqualified from participating in a study and/or unconsciously influence the results of the study using their GUILTY KNOWLEDGE. If you reveal the details of this study to another person we will also be forced to terminate your participation in this study.

Please sign this form in the space provided to indicate that you understand the instructions provided above.

_____

Participant Signature

_____

Printed Name

_____

Date

Appendix I

Participant Debriefing Statement II

Now that you have completed your role in our research, it is the desire of the entire project staff to take this opportunity to sincerely thank you for your help. Your work here may be more important than you realize.

The results of this study may include information which will provide federal agencies and police departments with a better understanding of how to change existing PDD examinations to accurately determine when an individual is being truthful.

If you participated in deceiving the PDD examiner, you are assured by the staff of this institute, that you in no way violated any rule or law. The deception was required for investigational purposes only.

Regardless of the role you played, it is our hope that you were made to feel as comfortable as possible throughout the study. If you do have concerns or questions regarding your participation, please make them known to the principal investigator, Dr. Andrew Dollins, and / or the DoD Polygraph Institute Director, Dr. William Yankee [Telephone number: (205) 848-3803].

Finally, it is VERY IMPORTANT that you DO NOT discuss the details of this study with anyone else. One of your friends, or a friend of a friend, may decide to participate in this or a similar study someday. If they know the details of the investigation process, they could be disqualified from participating in a study and/or unconsciously influence the results of the study using their GUILTY KNOWLEDGE.

Please sign this form in the space provided to indicate that you understand the instructions provided above.

_____

Participant Signature

_____

Printed Name

_____

Date

Appendix J

Power of ANOVA F-tests

Pn1-LnL: Power of 2(group) x 2(day) x 6(test) x 6(question) ANOVA F-tests ($p < .05$)

	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Group	0.0907	0.2545	0.5022	0.7449
Day	0.1448	0.4572	0.7927	0.9582
Group*Day	0.1448	0.4572	0.7927	0.9582
Test	0.1957	0.6965	0.9760	0.9998
Group*Test	0.1957	0.6965	0.9760	0.9998
Question	0.1957	0.6965	0.9760	0.9998
Group*Question	0.1957	0.6965	0.9760	0.9998
Day*Test	0.3797	0.9575	0.9999	0.9999
Group*Day*Test	0.3797	0.9575	0.9999	0.9999
Day*Question	0.3797	0.9575	0.9999	0.9999
Group*Day*Question	0.3797	0.9575	0.9999	0.9999
Test*Question	0.5838	0.9997	0.9999	0.9999
Group*Test*Question	0.5838	0.9997	0.9999	0.9999
Day*Test*Question	0.9300	0.9999	0.9999	0.9999
Group*Day*Test*Question	0.9300	0.9999	0.9999	0.9999

Pn2-LnL: Power of 2(group) x 2(day) x 6(test) x 6(question) ANOVA F-tests ($p < .05$)

	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Group	0.0907	0.2545	0.5022	0.7449
Day	0.1448	0.4572	0.7927	0.9582
Group*Day	0.1448	0.4572	0.7927	0.9582
Test	0.1957	0.6963	0.9760	0.9998
Group*Test	0.1957	0.6963	0.9760	0.9998
Question	0.1957	0.6963	0.9760	0.9998
Group*Question	0.1957	0.6963	0.9760	0.9998
Day*Test	0.3796	0.9575	0.9999	0.9999
Group*Day*Test	0.3796	0.9575	0.9999	0.9999
Day*Question	0.3796	0.9575	0.9999	0.9999
Group*Day*Question	0.3796	0.9575	0.9999	0.9999
Test*Question	0.5836	0.9997	0.9999	0.9999
Group*Test*Question	0.5836	0.9997	0.9999	0.9999
Day*Test*Question	0.9300	0.9999	0.9999	0.9999
Group*Day*Test*Question	0.9300	0.9999	0.9999	0.9999

GSR-LnL: Power of 2(group) x 2(day) x 6(test) x 6(question) ANOVA F-tests ($p < .05$)

	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Group	0.0857	0.2330	0.4620	0.7003
Day	0.1341	0.4196	0.7502	0.9386
Group*Day	0.1341	0.4196	0.7502	0.9386
Test	0.1790	0.6462	0.9608	0.9993
Group*Test	0.1790	0.6462	0.9608	0.9993
Question	0.1790	0.6462	0.9608	0.9993
Group*Question	0.1790	0.6462	0.9608	0.9993
Day*Test	0.3444	0.9355	0.9998	0.9999
Group*Day*Test	0.3444	0.9355	0.9998	0.9999
Day*Question	0.3444	0.9355	0.9998	0.9999
Group*Day*Question	0.3444	0.9355	0.9998	0.9999
Test*Question	0.5291	0.9990	0.9999	0.9999
Group*Test*Question	0.5291	0.9990	0.9999	0.9999
Day*Test*Question	0.8964	0.9999	0.9999	0.9999
Group*Day*Test*Question	0.8964	0.9999	0.9999	0.9999

GSR-Amp: Power of 2(group) x 2(day) x 6(test) x 6(question) ANOVA F-tests ($p < .05$)

	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Group	0.0857	0.2330	0.4620	0.7003
Day	0.1341	0.4196	0.7502	0.9386
Group*Day	0.1341	0.4196	0.7502	0.9386
Test	0.1789	0.6460	0.9607	0.9993
Group*Test	0.1789	0.6460	0.9607	0.9993
Question	0.1789	0.6460	0.9607	0.9993
Group*Question	0.1789	0.6460	0.9607	0.9993
Day*Test	0.3443	0.9354	0.9998	0.9999
Group*Day*Test	0.3443	0.9354	0.9998	0.9999
Day*Question	0.3443	0.9354	0.9998	0.9999
Group*Day*Question	0.3443	0.9354	0.9998	0.9999
Test*Question	0.5289	0.9990	0.9999	0.9999
Group*Test*Question	0.5289	0.9990	0.9999	0.9999
Day*Test*Question	0.8963	0.9999	0.9999	0.9999
Group*Day*Test*Question	0.8963	0.9999	0.9999	0.9999

GSR-Ltc: Power of 2(group) x 2(day) x 6(test) x 6(question) ANOVA F-tests ($p < .05$)

	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Group	0.0761	0.1897	0.3756	0.5923
Day	0.1132	0.3399	0.6428	0.8715
Group*Day	0.1132	0.3399	0.6428	0.8715
Test	0.1465	0.5292	0.9010	0.9945
Group*Test	0.1465	0.5292	0.9010	0.9945
Question	0.1465	0.5292	0.9010	0.9945
Group*Question	0.1465	0.5292	0.9010	0.9945
Day*Test	0.2730	0.8588	0.9978	0.9999
Group*Day*Test	0.2730	0.8588	0.9978	0.9999
Day*Question	0.2730	0.8588	0.9978	0.9999
Group*Day*Question	0.2730	0.8588	0.9978	0.9999
Test*Question	0.4125	0.9915	0.9999	0.9999
Group*Test*Question	0.4125	0.9915	0.9999	0.9999
Day*Test*Question	0.7888	0.9999	0.9999	0.9999
Group*Day*Test*Question	0.7888	0.9999	0.9999	0.9999

CRD-IBI: Power of 2(group) x 2(day) x 6(test) x 6(question) ANOVA F-tests ($p < .05$)

	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Group	0.0907	0.2545	0.5022	0.7449
Day	0.1448	0.4572	0.7927	0.9582
Group*Day	0.1448	0.4572	0.7927	0.9582
Test	0.1958	0.6966	0.9761	0.9998
Group*Test	0.1958	0.6966	0.9761	0.9998
Question	0.1958	0.6966	0.9761	0.9998
Group*Question	0.1958	0.6966	0.9761	0.9998
Day*Test	0.3798	0.9576	0.9999	0.9999
Group*Day*Test	0.3798	0.9576	0.9999	0.9999
Day*Question	0.3798	0.9576	0.9999	0.9999
Group*Day*Question	0.3798	0.9576	0.9999	0.9999
Test*Question	0.5839	0.9997	0.9999	0.9999
Group*Test*Question	0.5839	0.9997	0.9999	0.9999
Day*Test*Question	0.9301	0.9999	0.9999	0.9999
Group*Day*Test*Question	0.9301	0.9999	0.9999	0.9999

Pn1-LnL: Power of 2(day) x 6(test) x 6(question) ANOVA
F-tests ($p < .05$) of Deceptive and Non-deceptive Group Data

Group	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Deceptive				
Day	0.0896	0.2445	0.4823	0.7230
Test	0.1104	0.3691	0.7406	0.9520
Question	0.1104	0.3691	0.7406	0.9520
Day*Test	0.1908	0.6822	0.9723	0.9997
Day*Question	0.1908	0.6822	0.9723	0.9997
Test*Question	0.2709	0.9243	0.9999	0.9999
Day*Test*Question	0.5729	0.9996	0.9999	0.9999
Non-deceptive				
Day	0.0896	0.2445	0.4823	0.7230
Test	0.1104	0.3691	0.7406	0.9520
Question	0.1104	0.3691	0.7406	0.9520
Day*Test	0.1908	0.6822	0.9723	0.9997
Day*Question	0.1908	0.6822	0.9723	0.9997
Test*Question	0.2709	0.9243	0.9999	0.9999
Day*Test*Question	0.5729	0.9996	0.9999	0.9999

Pn2-LnL: Power of 2(day) x 6(test) x 6(question) ANOVA
F-tests ($p < .05$) of Deceptive and Non-deceptive Group Data

Group	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Deceptive				
Day	0.0894	0.2434	0.4801	0.7205
Test	0.1104	0.3689	0.7403	0.9519
Question	0.1104	0.3689	0.7403	0.9519
Day*Test	0.1907	0.6818	0.9723	0.9997
Day*Question	0.1907	0.6818	0.9723	0.9997
Test*Question	0.2708	0.9242	0.9999	0.9999
Day*Test*Question	0.5727	0.9996	0.9999	0.9999
Non-deceptive				
Day	0.0894	0.2434	0.4801	0.7205
Test	0.1103	0.3685	0.7397	0.9516
Question	0.1103	0.3685	0.7397	0.9516
Day*Test	0.1906	0.6812	0.9721	0.9997
Day*Question	0.1906	0.6812	0.9721	0.9997
Test*Question	0.2705	0.9239	0.9999	0.9999
Day*Test*Question	0.5722	0.9996	0.9999	0.9999

GSR-LnL: Power of 2(day) x 6(test) x 6(question) ANOVA
F-tests ($p < .05$) of Deceptive and Non-deceptive Group Data

Group	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Deceptive				
Day	0.0847	0.2213	0.4371	0.6703
Test	0.1031	0.3325	0.6874	0.9264
Question	0.1031	0.3325	0.6874	0.9264
Day*Test	0.1737	0.6282	0.9542	0.9990
Day*Question	0.1737	0.6282	0.9542	0.9990
Test*Question	0.2425	0.8875	0.9997	0.9999
Day*Test*Question	0.5164	0.9987	0.9999	0.9999
Non-deceptive				
Day	0.0848	0.2226	0.4400	0.6738
Test	0.1033	0.3335	0.6890	0.9273
Question	0.1033	0.3335	0.6890	0.9273
Day*Test	0.1742	0.6298	0.9548	0.9990
Day*Question	0.1742	0.6298	0.9548	0.9990
Test*Question	0.2430	0.8883	0.9997	0.9999
Day*Test*Question	0.5175	0.9987	0.9999	0.9999

GSR-Amp: Power of 2(day) x 6(test) x 6(question) ANOVA
F-tests ($p < .05$) of Deceptive and Non-deceptive Group Data

Group	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Deceptive				
Day	0.0847	0.2213	0.4371	0.6703
Test	0.1031	0.3322	0.6870	0.9262
Question	0.1031	0.3322	0.6870	0.9262
Day*Test	0.1736	0.6277	0.9540	0.9990
Day*Question	0.1736	0.6277	0.9540	0.9990
Test*Question	0.2424	0.8873	0.9997	0.9999
Day*Test*Question	0.5162	0.9987	0.9999	0.9999
Non-deceptive				
Day	0.0847	0.2213	0.4371	0.6703
Test	0.1032	0.3330	0.6883	0.9269
Question	0.1032	0.3330	0.6883	0.9269
Day*Test	0.1740	0.6290	0.9545	0.9990
Day*Question	0.1740	0.6290	0.9545	0.9990
Test*Question	0.2428	0.8880	0.9997	0.9999
Day*Test*Question	0.5171	0.9987	0.9999	0.9999

GSR-Ltc: Power of 2(day) x 6(test) x 6(question) ANOVA
F-tests ($p < .05$) of Deceptive and Non-deceptive Group Data

Group	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Deceptive				
Day	0.0759	0.1774	0.3453	0.5486
Test	0.0893	0.2600	0.5608	0.8385
Question	0.0893	0.2600	0.5608	0.8385
Day*Test	0.1412	0.5052	0.8839	0.9923
Day*Question	0.1412	0.5052	0.8839	0.9923
Test*Question	0.1890	0.7711	0.9958	0.9999
Day*Test*Question	0.3975	0.9892	0.9999	0.9999
Non-deceptive				
Day	0.0759	0.1792	0.3500	0.5556
Test	0.0895	0.2616	0.5640	0.8412
Question	0.0895	0.2616	0.5640	0.8412
Day*Test	0.1419	0.5082	0.8862	0.9926
Day*Question	0.1419	0.5082	0.8862	0.9926
Test*Question	0.1897	0.7733	0.9959	0.9999
Day*Test*Question	0.3993	0.9895	0.9999	0.9999

CRD-IBI: Power of 2(day) x 6(test) x 6(question) ANOVA
F-tests ($p < .05$) of Deceptive and Non-deceptive Group Data

Group	Effect Size			
F-test Effect	0.10	0.20	0.30	0.40
Deceptive				
Day	0.0896	0.2445	0.4823	0.7230
Test	0.1105	0.3693	0.7409	0.9521
Question	0.1105	0.3693	0.7409	0.9521
Day*Test	0.1910	0.6825	0.9724	0.9997
Day*Question	0.1910	0.6825	0.9724	0.9997
Test*Question	0.2710	0.9244	0.9999	0.9999
Day*Test*Question	0.5731	0.9996	0.9999	0.9999
Non-deceptive				
Day	0.0896	0.2445	0.4823	0.7230
Test	0.1104	0.3691	0.7406	0.9520
Question	0.1104	0.3691	0.7406	0.9520
Day*Test	0.1908	0.6822	0.9723	0.9997
Day*Question	0.1908	0.6822	0.9723	0.9997
Test*Question	0.2709	0.9243	0.9999	0.9999
Day*Test*Question	0.5729	0.9996	0.9999	0.9999